

Archive-It Partner Meeting
Thursday, October 18, 2007
Hawthorne Room of the Golden Gate Club in the Presidio of San Francisco
Partner Presentations I, Collection Development
10:00-11:00 a.m.

Indiana University
How Archive-It fits into Indiana University Libraries' Access and Preservation Efforts

Introduction:

Many thanks to Archive-It for organizing this meeting. Being here in the Presidio gives us tangible evidence of change, and various paces for change, as we share our experiences with Archive-It. I'm Andrea Singer, Foreign Documents Librarian and Librarian for India and Tibetan Studies at Indiana University Libraries in Bloomington, Indiana. I'm delighted to talk in the "collection development" session of the program, because the evolution of Indiana's participation has seated our use of Archive-it directly in collection managers' hands. (At this point, since we're in California, and in the Presidio, imagine a movie: three people sit at their workstations, developing collections. Technical advances and support are based here in San Francisco. Our crawls are like unmanned space flight into the web world, capturing known and unknown quantities.)

I'll wrap up with more on this imagery, but first will describe the history of Archive-It at Indiana, and note some access tools we've developed. Then in a more conversational mode I'll turn to the topic of the talk: how Archive-It fits into our access and preservation efforts. I hope we'll have lots of time for questions and discussion.

History at Indiana:

We've been using Archive-It since September 2005, when we began as participants in the RLG beta test. Our dean at the time, Suzanne Thorin, now at Syracuse, supported our testing from the top of the organization. Like much at Indiana, the work took place in a committee or working group. This one had participation from Cataloging, the University Archives (which reports to the Libraries), and Collection Development staff. Technology staff was not involved.

After a test period in which several collection managers, including the director of the Indiana University Archives, experimented with URLs related to their areas, development began in earnest. In spring 2006 we subscribed, and began work on three collections.

Issues we considered in choosing the collections:

- Need to archive and preserve "born digital" content for our users
- Web sites of local interest

- Level of specificity Archive-It provides (web sites over time, not documents over time, for example)
- Cooperative collection development: where can we contribute?

The three collections chosen are described in the brochure I'm distributing: *Web Archives at Indiana University: Collections in the Indiana University Libraries*. Two are of great local interest: Indiana University Web Sites, and Indiana: State and Local Documents. I'm the collection manager for the Foreign Government Statistical Web Site – one that is not so obvious a choice. (Statistical yearbooks have been an important and much-sought genre in our collecting for many years. Here is a sample of our collection brochure for the Foreign Government Information collection, which places the Archive-It collection in a context of other resources.)

The three collection managers were assisted within the committee context by catalogers, who developed a template for adding records for the overall Web Archives collection and the individual collections to our online catalog, IUCAT. They also contributed to our discussion of metadata, and decisions to take a lean approach for now. They treated all the collections consistently in the catalog.

In the Libraries' web site we also treat the Archive-It collections in parallel to print or mixed collections:

(Show the following urls of an overview, a sample collection link, and a sample finding aid:)

<http://www.libraries.iub.edu/index.php?pageId=4302>

<http://www.archive-it.org/collections/219>

<http://www.libraries.iub.edu/index.php?pageId=5024>

The division of labor for staff currently is that a librarian in the University Archives and two collection managers in Government Information are the managers choosing seeds, reviewing and monitoring crawls, and testing new features. (In general, web content development is also distributed to collection managers at Indiana, so the Archive-it role is not an unusual direction.)

Cecile Jagodzinski, the Director of Collection Development, is the central information and budget manager for the entire project. She oversees the crawl budget, and decides when we can consider new initiatives.

The Committee is somewhat active, and is supported by the administrative assistant to the Director of Collection Development, who is the most tech-savvy member of the group. (When Molly Bragg sends upgrade messages, Charla Lancaster is normally the person assigned to test implications, though Dina Kellams of the University Archives is always eager to test anything that provides more specific targeting.)

After some initial tinkering with timing of crawls, and some limiting, we have generally chosen a low maintenance approach to managing URLs before and after a crawl.

Users:

The committee hosted an internal information session for librarians and library staff, and the Libraries' Director of Communications wrote short articles on our collections for our web site and University publications targeted to an external audience.

We haven't analyzed use. Public Services and internal staff have searched for specific "lost" items with success, especially on the archives side. (IU has a new President, and new administrators at top levels, so we are alert to the benefits of archived web sites as administrations change in both the Archive and the State and Local collections.)

Where we are now:

We view our participation as a project that may take some time to evaluate.

Other aspects of our digital preservation/access mélange are also in development. The Libraries use D-Space for archiving scholarly works of faculty and librarians through IUScholarWORKS, which we have ruled unavailable for administrative use. The archive is maintained by the IU Libraries and the IU Digital Library Program. Here is the link for IU ScholarWORKS:

<http://scholarworks.iu.edu>

Recently the University decided to fund a separate D-Space environment for born digital and other documentary access and preservation uses including administrative ones. The University Archives will be a key investor in this effort, which will enable them to design a space at the document level. (Class bulletins, for example – in a series...)

Other digital preservation initiatives the Libraries are supporting include membership in LOCKSS and Portico.

The committee is seeking additional collection managers who are eager to use Archive-it for their collections. Each conversation about the possibility leads to higher visibility for the product.

An interesting example is that in a conversation with our Folklore librarian I learned that there is great interest among graduate students in State Folklore and Folklife initiatives, which might have ephemeral lives on the web. (Public folklore programs are possible sources of job opportunities for the professional folklorists Indiana trains.) The Folklore librarian was very eager to explore the use partners in State archives are making of Archive-It. (In that conversation I think we gained an Archive-It user, but not a Web Archives at IU collection developer.)

Conclusion:

As I conclude I'll return to the NASA movie imagery. The people at the content computers at IU are wearing many hats and usually responsible for several collections in combinations of administrative, reference and public service roles. That means that developments managed centrally by Archive-It are highly valued since we need a low maintenance, reliable product for the long term. We are delighted to share with other partners. "Hello San Francisco", instead of "Hello Houston".

AS 10/23/07